



Development of an online *Nigella sativa* inspection system equipped with machine vision technology and artificial neural networks

✉ Saman ALVANDI, ✉ Seyed Saeid MOHTASEBI*, ✉ Mahmoud OMID and ✉ Mohammad HOSSEINPOUR-ZARNAQ

Department of Agricultural Machinery Engineering, Faculty of Agriculture, University of Tehran, Karaj, Iran.

*Correspondence should be addressed to Seyed Saeid Mohtasebi: mohhtaseb@ut.ac.ir

Abstract

Aim of study: *Nigella sativa* L. seeds usually are mixed with impurities, which affect its quality and influences consumer acceptance in both raw seeds and the oil market. In this study, an intelligent system based on the combination of machine vision (MV) and artificial neural networks (ANN) was developed to classify and clean *N. sativa* seeds and its impurities.

Area of study: Iran, Kurdistan province.

Material and methods: For accurate detections we developed a robust image processing algorithm including image acquisition, image enhancement, segmentation, and feature extraction steps. Correlation-based Feature Selection method was used to select the superior features. Three methods of linear discriminant analysis, support vector machines, and ANN were used to classify the data.

Main results: The statistical indices of sensitivity, specificity, and accuracy for *N. sativa* in the online phase were 90%, 98.93%, and 97.04%, respectively. The average of these measurements for the impurities class was 95.57%, 96.89%, and 96.58%, respectively.

Research highlights: The results demonstrated the feasibility of suggested machine learning and image processing approaches in the real-time cleaning of *N. sativa*. The image acquisition and processing process, including selection of the best lighting methods to reduce the shadows, noise elimination and segmentation, provided precise results. The final results indicated the effectiveness of proposed machine learning algorithm in feature extraction, feature dimensionality reduction, and classification approaches. This methodology can be recommended for detection, classification and automatic cleaning of other similar seeds.

Additional key words: sorting; cleaning; image processing; artificial intelligence.

Abbreviation used: ANN (artificial neural networks); BP (back-propagation); GA (genetic algorithms); IBT (inverted binary thresholding); ICA (independent component analysis); IZT (inverted to zero thresholding); L-BFGS (limited memory Broyden–Fletcher–Goldfarb–Shanno); LDA (linear discriminant analysis); MLP (multilayer perceptron); PCA (principal components analysis); SVM (support vector machines).

Citation: Alvandi, S; Mohtasebi, SS; Omid, M; Hosseinpour-Zarnaq, M (2023). Development of an online *Nigella sativa* inspection system equipped with machine vision technology and artificial neural networks. Spanish Journal of Agricultural Research, Volume 21, Issue 2, e0206. <https://doi.org/10.5424/sjar/2023212-19317>.

Received: 20 Feb 2022. **Accepted:** 08 May 2023.

Copyright © 2023 CSIC. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License

Funding agencies/institutions	Project / Grant
Department of Agricultural Machinery Engineering, University of Tehran, Iran	7187730

Competing interests: The authors have declared that no competing interests exist.

Introduction

Nigella sativa L., as a popular medicinal plant, has one hundred different chemicals including essential acids (Majdalawieh & Fayyad, 2016). Although *N. sativa* does not have a significant weight in the global market, its important medicinal properties increase its usage constantly (Gharby et al., 2015). The raw seed and the extracted oil are broadly used in food and pharmaceutical applications. Separation, cleaning, and detection of contaminants are the most essential issues in grains post-harvest processing. Due to the small size of *N. sativa* and its impurities, the separation process has some challenges (Copeland & McDonald, 2012). Machine vision has been one of the most powerful, inexpensive, non-destructive, accurate, and reliable measurement tools in grain technology for the past two decades (Vithu & Moses, 2016). Advances in image processing and machine learning techniques have greatly improved this tool, so this technology with new controllers can classify grains with similar physical characteristics, eliminating the disadvantages of mechanical cleaning systems such as low accuracy, grain damage, and contamination of the product with other impurities such as oil vapor (Rad et al., 2011). Seeds have a broad distribution of visual characteristics. So various morphological, color, and texture properties are needed for their distinction and grading. Also, advanced analytics approaches such as artificial intelligence techniques are required for the analysis of these properties (Orlandi et al., 2018; Qadri et al., 2021).

Machine vision have been used for separation of grains, diseased grains, moldy grains and impurities. Rad et al. (2011) classified five rice cultivars by a back-propagation (BP) artificial neural network (ANN) with 97% accuracy. Determining the number of hidden layers and neurons is very important to achieve a suitable model for mining the data. However, it is a difficult task and requires trial and error. Han et al. (2014) reported that the performance of a support vector machine (SVM) with genetic algorithms (GA) for the selection of superior features from wheat seed images achieved better results than linear discriminant analysis (LDA) and principal components analysis (PCA). The optimum population size of the GA was 24. The results showed that GA-SVM successfully selected and classified the features. Two features, namely, red in RGB space and core area, were sufficient for the online classifying of wheat grains. Ebrahimi et al. (2014) presented an algorithm based on ANN with independent component analysis (ICA) to classify the wheat seeds and determine the degree of purity. In their imaging system, the seeds were separated from each other by hand to avoid contact with each other. Their algorithm was able to accurately detect wheat grains, wheat cultivars, and weed seeds with an accuracy of 96.25%, 87.5%, and 77.22%, respectively. In another study, Liu et al. (2015) showed that using only texture features were not sufficient to show differences in damage, breaks, and partially incomplete skin of soybeans.

Shrestha et al. (2016) measured the contrast, correlation, energy, and homogeneity of wheat grains in the horizontal and vertical directions to extract texture characteristics. To identify three Iranian rice species, Fayyazi et al. (2017) reports indicating the possibility of mixing different varieties on the market. Applying machine vision techniques to classify rice varieties is a method which can increase the accuracy of classification process in real applications. In this study, several morphological and textural features of rice seeds' images were examined to evaluate their efficacy in identification of three Iranian rice varieties (Tarom, Fajr, Shiroodi) extracted morphological and texture features and used the PCA method to select and rank the most important ones; the rice classification was performed using multilayer perceptron (MLP) neural networks. For the detection of discolored peppercorn grains, Kanjanawanishkul et al. (2018) used two processors, a PC-based controller processed images and detected discolored seeds, and an Arduino controller started and stopped the rotation of the circular plate containing the seeds and controlled the pneumatic valves. This machine accurately classified 90.9% of the input grains.

The impurity affects the oil extraction process of *N. sativa* and influence consumer acceptance in both raw seeds and oil market. Accordingly, the automation of the cleaning process is essential for performing high-quality products. The mechanical separation of *N. sativa* from impurities based on the pneumatic, friction, and sieve mechanisms has not been implemented due to their similar physical properties (geometric characteristics, static friction coefficient, and terminal velocity). Therefore, in this study, an intelligent cleaning system based on novel non-destructive image processing techniques and machine learning algorithms was designed and developed. The device automatically detected impurities by combining the machine vision system and the ANN model.

Material and methods

Figure 1 shows a scheme of the developed automatic system for sorting impurities from *N. sativa*. Python platform was used to develop the algorithms. Further information is presented in the following subsections.

The *N. sativa* seeds were collected directly from Kurdistan province, Iran. The grains were transferred to the Department of Agricultural Machinery Engineering, University of Tehran. The product contained several impurities such as weed seeds and soil particles (Fig. 2).

Components of the designed system

This device consists of five main parts: a conveyor belt, an imaging box, an electronic part, a separator, and a personal computer (Fig. 1). In this research, a white con-

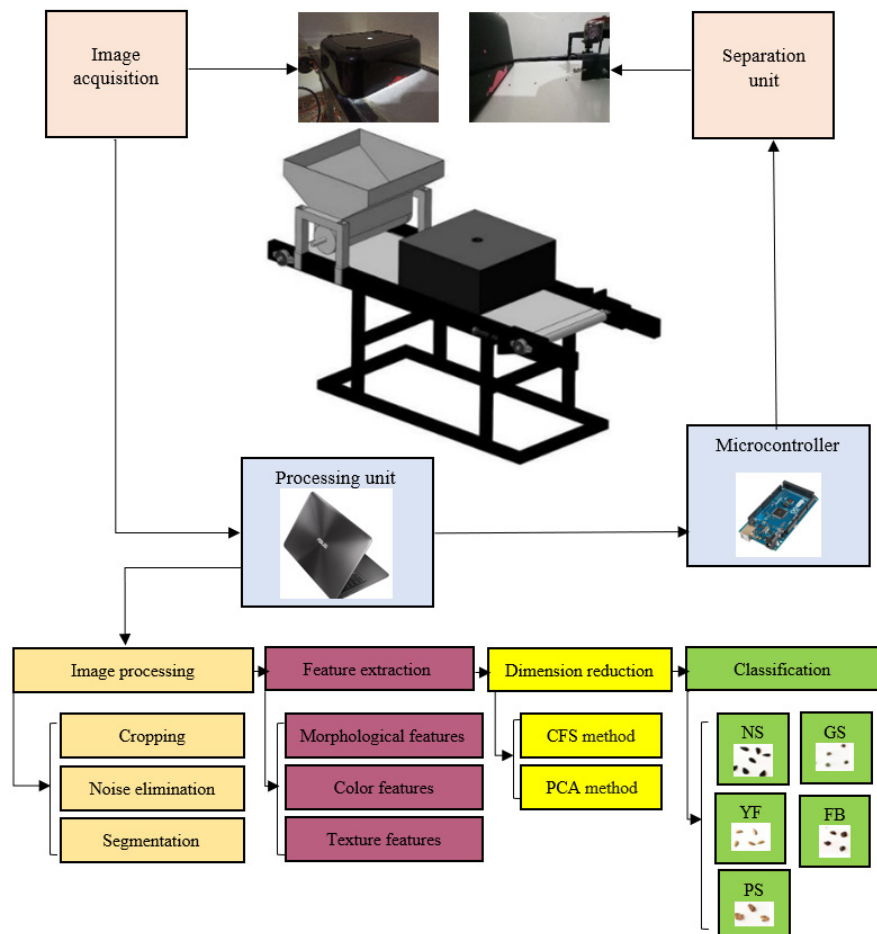


Figure 1. The proposed system for separating *Nigella sativa* and its impurities. CFS: correlation-based feature selection, PCA: principal component analysis, NS: *Nigella sativa*, FB: brown seed, YF: yellow flax, PS: soil particles, GS: grey seed.

veyor belt was used to achieve better segmentation. The conveyor belt transferred the product under the camera and the separator. The length, width, and thickness of the belt were selected as 1.5 m, 20 cm, and 1 mm, respectively. To move the conveyor belt with a constant speed and with high torque, a DC motor (12 volts with a maximum speed of 150 rpm) with two conical gears was selected. Motor speed can be easily adjusted by controlling the voltage volume. Lighting is one of the most important parts of a machine vision system. To control the light and receive homogenous light, a black rectangular box with four LEDs was used to create uniform lighting conditions and maximize the reduction of shadows. A 64-megapixel mobile camera with the specifications of the Xiaomi Note 8 Pro was used for capturing. The images were transferred wirelessly to the computer for further processing through the IP Webcam software installed on the mobile phone. Image enhancement, segmentation, noise cancellation, feature extraction, superior feature selection, and classification were performed on the computer (Core i7-5500U CPU; 2.4 GHz; 8 GB RAM; Windows 10, 64 bits). The commands were sent to the separator via an ATmega2560 Arduino board.

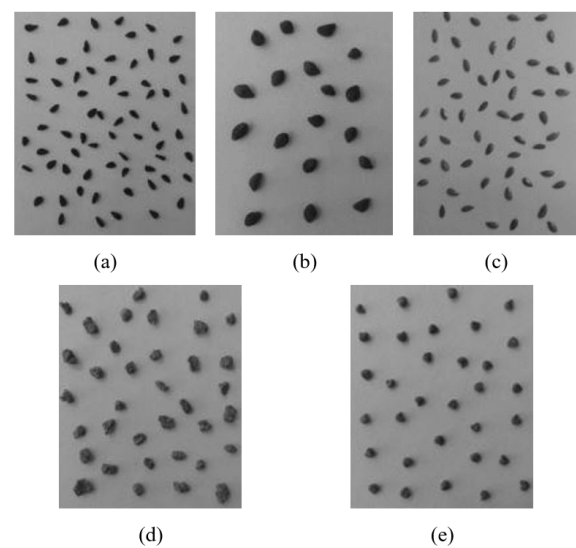


Figure 2. Image capture technique using the proposed hybrid lighting system: a) *Nigella sativa*, b) brown seed (*Convolvulus arvensis*), c) yellow flax, d) soil particles, and e) grey seed (*Chenopodium album*).

Table 1. Morphology (shape) features (F-1 to F-9) computed for *Nigella sativa* and its impurities.

Definition	Formula	Feature No.
Dimension	Area	F-1
Dimension	Perimeter	F-2
Equivalent diameter	$\sqrt{\frac{4(Area)}{\pi}}$	F-3
Circularity ratio	$\frac{4\pi (Area)}{(Perimeter)^2}$	F-4
Dimension	$\frac{Perimeter}{Area}$	F-5
Dimension	$\frac{area}{\text{Minor diameter}}$	F-6
Aspect ratio	$\frac{\text{Major diameter}}{\text{Minor diameter}} = \frac{H}{W}$	F-7
Extent	$\frac{Area}{WH}$	F-8
Solidity	$\frac{Area}{\text{Hull_area}}$	F-9

A servomotor model MG90s separated the impurities from *N. sativa* seeds, by deflecting the grains from their original path. It has a lever that moves a flexible object in the grain flow path, and can be rotated from zero degrees to the 180-degree angle. It was located at the angle of 45 degrees for *N. sativa* and the angle of 135 degrees for impurities.

The area of the image where the grain is located was cropped after receiving it. The cropping operation reduces the volume of calculations and speeds up the grading process. To remove noise and fill cavities, the dilation and erosion operation were applied.

Image segmentation

Images in RGB color space were decomposed into three channels (R, G and B), and a binary threshold was applied to the image to separate the grains from the background. In the thresholding method, the whole image was divided into windows with dimensions of 3×3, 5×5, and 7×7. Two thresholding methods, namely, inverted binary thresholding (IBT) and inverted to zero thresholding (IZT) were utilized for segmentation. The threshold value was determined by computing the mean and standard deviation of the brightness of pixels for each category. So all are-

Table 2. Color features (F-10 to F-45) computed for the automatic separating of *Nigella sativa*. See equations in the text.

Color component								
R	G	B	R/R+G+B	G/R+G+B	B/R+G+B	R-B	R-G	G-B
Eq [1]								
F-10	F-11	F-12	F-13	F-14	F-15	F-16	F-17	F-18
Eq [2]								
F-19	F-20	F-21	F-22	F-23	F-24	F-25	F-26	F-27
Eq [3]								
F-28	F-29	F-30	F-31	F-32	F-33	F-34	F-35	F-36
Eq [4]								
F-37	F-38	F-39	F-40	F-41	F-42	F-43	F-44	F-45

Table 3. Texture features (F-46 to F-106) extracted from *Nigella sativa* and its impurities.

Distances and angles ^a	Contrast	Energy	Homogeneity	Correlation	Dissimilarity	Entropy
[1, 0]	F-46	F-58	F-70	F-82	F-94	F-106
[1, 45]	F-47	F-59	F-71	F-83	F-95	
[1, 90]	F-48	F-60	F-72	F-84	F-96	
[1, 135]	F-49	F-61	F-73	F-85	F-97	
[2, 0]	F-50	F-62	F-74	F-86	F-98	
[2, 45]	F-51	F-63	F-75	F-87	F-99	
[2, 90]	F-52	F-64	F-76	F-88	F-100	
[2, 135]	F-53	F-65	F-77	F-89	F-101	
[3, 0]	F-54	F-66	F-78	F-90	F-102	
[3, 45]	F-55	F-67	F-79	F-91	F-103	
[3, 90]	F-56	F-68	F-80	F-92	F-104	
[3, 135]	F-57	F-69	F-81	F-93	F-105	

^aDistances in pixels and angles in degrees

as of the same intensity were considered as one object. In this way, the seed and the background could be completely separated (Fig. 3).

Feature extraction

Color, texture, size, and shape are the most important visual features of an object in the image (Nouri-Ahmadabadi et al., 2017). A total of 106 features (9 morphology features, 36 color features, and 61 texture features) were extracted from each sample. *N. sativa* seeds and their impurities differ slightly in color, so the extraction of all morphological, color, and texture properties is necessary for accurate classification.

The extracted morphology and color features with their formulas are listed in Tables 1 and 2, respectively. Due to the small color difference between some impurities and *N. sativa*, the four Bao et al.'s (2019) statistical descriptors were computed:

$$\text{Mean} \left[\mu = \frac{\sum_{x=0}^{z-1} h(x)}{z} \right] \quad (1)$$

$$\text{Variance} \left[\sigma = \frac{\sum_{x=0}^{z-1} (h(x)-\mu)^2}{z} \right] \quad (2)$$

$$\text{Skewness} \left[S = \frac{\sum_{x=0}^{z-1} (h(x)-\mu)^3}{z\sigma^3} \right] \quad (3)$$

$$\text{Kurtosis} \left[K = \frac{\sum_{x=0}^{z-1} (h(x)-\mu)^4}{z\sigma^4} \right] \quad (4)$$

where $h(x)$ represents the grey level of pixels in the image with a pixel position of x , x can take any value between 1

and $z = m \times n$, m and n being the number of rows and columns of the image matrix, respectively.

For texture classification, the features used were roughness, smoothness, or regularity and irregularity of the grains surface (Nouri-Ahmadabadi et al., 2017; Bao et al., 2019). In some cases, the images of impurities and *N. sa-*

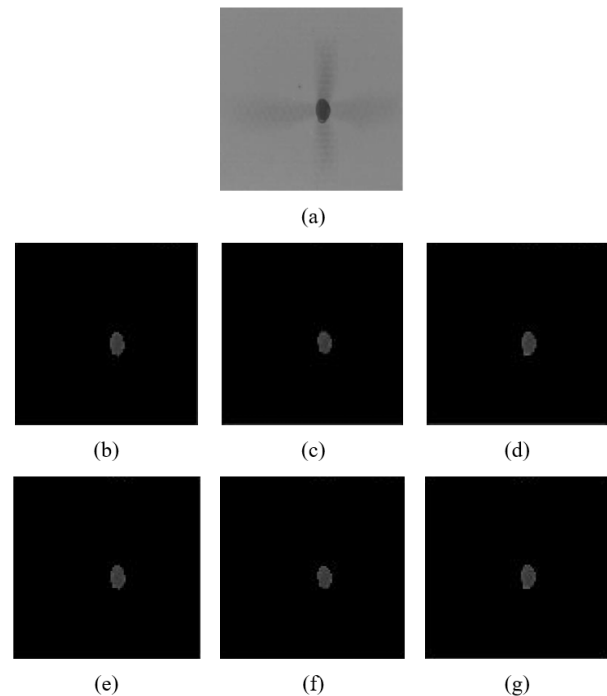


Figure 3. Image segmentation using the inverted binary thresholding (IBT) and inverted to zero thresholding (IZT) methods. a) Initial image, b) IBT on R channel, c) IBT on G channel, d) IBT on B channel, e) IZT on R channel, f) IZT on G channel, g) IZT on B channel.

Table 4. Selected features for classification of *Nigella sativa* and its impurities using correlation-based feature selection (CFS) analysis.

Morphological features	Color features	Texture features
F-5	F-10	F-22
F-7	F-11	F-24
F-8	F-12	F-27
F-9	F-13	F-28
	F-14	F-31
	F-15	F-33
	F-16	F-38
	F-18	F-41
	F-21	
		F-49
		F-83
		F-103
		F-106

tiva had similar color and shape, but the areas with different color intensities and greyscale levels created different tissue structures and could distinguish each class. Texture features including contrast, energy, homogeneity, correlation, dissimilarity, and entropy were extracted (Table 3) using the Grey level co-occurrence matrix according to Eq. (5) (Mollazade, 2017):

$$C_{\Delta x, \Delta y}(i, j) = \sum_{x=1}^n \sum_{y=1}^m \begin{cases} 1, & \text{if } I(xy) = i \text{ and } I(x + \Delta x, y + \Delta y) = j \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where C is the grey level co-occurrence matrix; m and n are the values related to the brightness; i and j are the image sizes; x and y are the spatial coordinates; and Δx and Δy are the deviation parameters. In this study, the distance parameter was chosen as 1, 2, and 3 pixels, and the angle value was selected as 0, 45, 90, and 135 degrees (Table 3).

Feature selection

A total of 100 samples for each class of *N. sativa* and impurities (brown seed, yellow flax, soil particles, and grey seed) were selected and 106 features were extracted from each image. The dataset normalized before feature selection and classification. In this research, the Z-score normalization method, as the most common technique, was utilized. The performance of the developed algorithm depends on the obtained features. The superior and efficient features increase the performance of the proposed algorithm and reduce the execution time. In data mining, there are many ways to select the superior features, such as sequential forward selection (SFS), factor analysis, correlation-based feature selection (CFS), genetic algorithm, and sensitivity analysis (Mollazade et al., 2012a; Valiente-Gonzalez et al., 2014). In this study, the CFS method was used to select important and efficient features from the feature set. The method was implemented in WEKA software version 3.9.

PCA is a simple and efficient linear analysis method. The main purpose of PCA is to identify patterns in the data. PCA intends to identify the correlations between variables. In this method, a large number of independent variables can be replaced with a limited number of new dependent variables called principal components. PCA reduces the dimension of the dataset (X) in the following five main steps (Jolliffe & Cadima, 2016; Shen et al., 2020):

- Decentralization of the values of features (value of each feature minus its average).
- Computing the covariance matrix for X.
- Eigenvalues and eigenvectors of the covariance matrix are obtained by eigenvalue decomposition or singular value decomposition.
- Sorting the eigenvalues from large to small, selecting the k largest eigenvectors, and then taking the corresponding k eigenvectors as column vectors to form the eigenvector matrix P.
- Transforming the data into the new space constructed by the k eigenvectors, i.e., $Y = PX$.

In this research, the CFS method reduced the dimension of extracted features and PCA reduced the size of top attributes while retaining all of the data information (Mollazade et al., 2012b). In this research, the scikit-learn open-source library was utilized for the implementation of machine learning algorithms. Also, 75% of the data were used during the model building step including training and validation (375 samples), and the remaining data (125 samples) were selected as test set for final evaluation.

Classification

— **Linear discriminant analysis (LDA).** LDA was used to classify *N. sativa* and its impurities. LDA is a statistical method utilized in machine learning and pattern recognition to find the linear combination of properties that best distinguish two or more classes of objects. LDA maximizes the ratio of between-group variance and within-group variance. When the value of between-group ratio is at its maximum,

Table 5. Confusion matrix representing the separation performance of *Nigella sativa* and its impurities in online mode.

	Class	Predicted				
		NS	FB	YF	GS	PS
Actual ^a	NS	45	0	0	5	0
	FB	2	48	0	0	0
	YF	0	0	50	0	0
	GS	0	6	0	44	0
	PS	0	5	0	2	43

^a Actual class samples (see Fig. 1) in the confusion matrix.

then the samples within each group have the smallest possible scatter (Stanimirova et al., 2013). The purpose of this method is mapping the features, so the different classes are further separated based on the combination of raw input features. To make a prediction, LDA estimates the probability that a new set of inputs belongs to every class. The output class is the one that has the highest probability.

— **Support vector machines (SVM).** Support vector machines are commonly used as one of the classification methods. SVM showed very good performance in various classification problems. The basis of SVM is the linear classification of data. SVM maps the input vector to a higher dimensional space using a kernel function and then finds a hyperplane in N-dimensional space (N is the number of features) that classifies the data. To separate the two classes of data points, there are many possibilities. A plane that has the maximum distance between data points of classes (maximum margin) is selected as a separating hyperplane. The following steps were implemented to select the most appropriate SVM model:

Several SVM models were developed with different kernel functions (linear, quadratic, polynomial, and radial basis), and SVM with a polynomial kernel was selected as the superior. Finally, the best model was selected based on the highest accuracy in the test phase while having the minimum number of support vectors. SVM with the minimum number of support vectors will increase processing speed and make the algorithm more suitable for use in the online mode.

Table 6. Performance measurements for online separation of *Nigella sativa* and its impurities.

Class ^a	Sensitivity (%)	Specificity (%)	Accuracy (%)
NS	90	98.93	97.04
FB	96	94.30	94.65
YF	100	100	100
GS	86.27	96.87	94.65
PS	100	96.39	97.04

^a See classes in Fig. 1.

— **Artificial neural network (ANN).** In this work, multilayer perceptron (MLP) neural networks were used for predicting the outputs. MLP model has three main layers: the input layer, the hidden layers, and the output layer. The input layer acts as a receiver and contains the input features to be fed into the neural model. Hidden layers compute the data undergoing iterations to reduce the overlap between patterns. The final layer predicts the output. The number of neurons in the output layer depends on the application. If the data were complex, higher-order features and a suitable ANN structure must be designed. To find the optimum structure of the ANN, the following steps were performed.

Various structures were evaluated in terms of the number of neurons and hidden layers. Thus, the network was evaluated in two modes, a single hidden layer, and a double hidden layer. The number of neurons with a single hidden layer was selected as 4, 8, 10... 16, and the number of neurons in the two hidden layers were selected as (4, 4), (8, 8), (10, 10), ..., (20,20), (6,8) and (8,10). In this research, two different activation functions were tested in the hidden layers including hyperbolic tangent and logistic activation functions. Also, the network was tested with three powerful optimizers, namely, adaptive moment estimation (Adam), stochastic gradient descent (SGD), and limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS). The Adam solver performs very well in large datasets (with thousands of training samples) in terms of training time and validation score. However, for small datasets, L-BFGS can converge faster. In the last step, the structures were repeated several times and the accuracy parameter was used to evaluate the results. The tests showed that logistic (sigmoid) activation functions and the Adam optimizer enriched the results. Finally, the accuracy index of LDA, SVM, and ANN were compared and the best classifier was selected for the online separation of *N. sativa*.

Performance evaluation

To evaluate the performance of the proposed method, three statistical indicators were used in the classification procedure. These parameters are sensitivity, specificity

and accuracy. These statistical parameters were calculated from the following equations:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (7)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

where TP is the impurity correctly detected in the impurity class; TN is the *N. sativa* correctly detected; FP is *N. sativa* that has been misdiagnosed as impurity; and FN is the impurity that has been misidentified as *N. sativa*.

Results and discussion

An Arduino board was used to send the commands to separator, and the Python programming language was used to develop the image processing algorithm and ANN classification. The flexibility of Python programming and the power of Arduino hardware provided a high-quality platform for the implementation of grain automatic cleaning projects.

The proposed lighting system was very effective in reducing object shadows which is one of the main problems in previous research (Quan et al., 2018). There is also no need for shadow removal steps in image processing algorithms, which may have influenced on the results of other investigations (Mollazade et al., 2012a; Pearson et al., 2012; Teimouri et al., 2014). As shown in Fig. 2, the shadow around the seeds was greatly reduced. In the capturing of yellow flax, some light was reflected due to the smooth surface of the grain; this reflection caused that some areas of yellow flax were considered as the background in the segmentation. The images were segmented in RGB color space with thresholding on R, G and B channels. Two thresholding methods, IBT and IZT, with a constant threshold value of 90, were utilized. Figure 3 indicates that the best result for both approaches were related to thresholding on channel B. IZT is one of the most popular segmentation methods and other researchers have utilized this approach in similar studies (Rad et al., 2011).

At the feature extraction and selection stages, CFS selected 25 features from a total of 106 features as the superior features. The results are shown in Table 4. The CFS selected four features from morphology, 17 features from color, and four features from texture data. The CFS has also been used in similar investigations for feature selection (Banakar et al., 2017). Selecting the optimal number of PCs in PCA is important to compress the size of the features. The first five components were selected as optimal

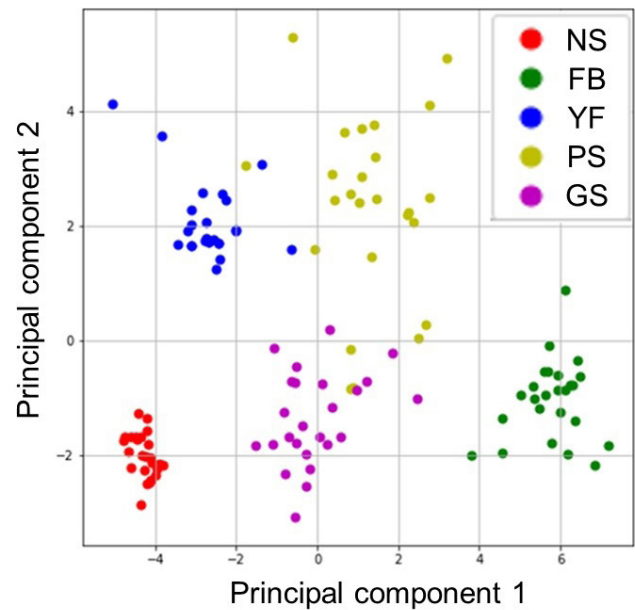


Figure 4. Visitation of *Nigella sativa* and its impurities data using the first two components of PCA. NS: *Nigella sativa*, FB: brown seed, YF: yellow flax, PS: soil particles, GS: grey seed.

PCs for keeping more than 95% of the input information. Therefore, the execution of machine learning algorithms could accelerate. Figure 4 shows a two-dimensional data map. It is indicated that the classes are successfully separated using only two PCs. Figure 4 shows that the *N. sativa* (red dots) was well distinguished from its impurities. However, the separation of soil particles from other impurities was relatively low. This may be due to the uniform color and texture of the *N. sativa* compared to other classes.

The first five PCs were given as inputs to the classifiers (LDA, SVM and ANN). The accuracy index in the offline mode for LDA, SVM and ANN were 71.8%, 88%, and 100%, respectively. According to the results, ANNs showed better performance than the other models. The LDA modeling provides an initial features transformation to compute the simple univariate variables that give maximum group separation in the transformed data (Wakholi et al., 2018). This approach could be precisely used in linear solutions. Also, SVM performance is attributable to the use of the kernel function, which maps the information to a higher dimensional space, which makes the SVM a feasible choice on non-linear nature of problem. SVM requires a lower number of input data as compared to ANN due to the size of the structure. The ANN structure employs multi-layer connection and activation for dealing with non-linear problems and computing complex patterns for describing the data. ANN could identify the seeds and quality using machine vision systems (Granitto et al., 2005; Rahman & Cho, 2016; Madhavan et al., 2020).

Therefore, ANN was used as a classifier in the online phase. These results are in line with Teimouri et al. (2018).

The accuracy parameter was calculated separately for the training data and test data to avoid over-fitting. If the ANN performance is very good for the training data and poor for the test data, this means that overfitting occurs. The best structure of ANN was 5-16-5, with an accuracy of 100% for offline test data. The same ANN structure was used for the online mode and the results are as follows: To evaluate the developed device's performance, 50 new grains of *N. sativa* samples and each of the impurities were randomly selected and evaluated. The confusion matrix for *N. sativa* and its impurities is shown in Table 5. Compared to results of offline mode, the errors of correct detection rates were increased in online mode, due to the movement of the grains and the very small size of the *N. sativa* and its impurities. The sensitivity, accuracy, and specificity of each class are shown in Table 6. The proposed model detects the impurities (other seeds) with high accuracy. This proves that the lightning structure could effectively reduce the shadows and distinguish the objects in the image. Results of ANN showed that the proposed model precisely detected the *N. sativa* class. The accuracy of ANN associated with feature selection for *N. sativa* samples was 97.04% in online mode.

The ANN model was successfully utilized for grading of seeds. Rad et al. (2011) classified five rice cultivars by ANN with 97% accuracy. Ebrahimi et al. (2014) detected wheat grains, cultivars and seeds using machine vision techniques and ANN algorithms with an accuracy of 96.25%, 87.5%, and 77.22%, respectively. Qadri et al. (2021) classified eight canola seed varieties using a multi-featured machine vision approach, including first-order histogram, second-order statistical texture, binary and spectral features from the bulk images of canola seeds; the accuracy of the ANN classifier for the eight varieties was 95-98%.

Due to the long response time of the servomotor, the distance between the samples on the conveyor was increased; therefore, the limiting factors of the system performance were related to the hardware components of the separator. Similar results were obtained in the fabrication of a machine for peeled pistachios with machine vision by Nouri-Ahmadabadiet al. (2017).

As a result of the first detection of *N. sativa* contaminations by ANN, the prediction accuracy could also constitute an objective tool for non-destructive and rapid quality evaluation with the preparation of automation of the cleaning process to remove the impurities. Moreover, the proposed method could lead to the development of automated systems for real-time cleaning of other similar products.

In summary, the present study showed that the combination of machine vision and artificial neural networks (ANN) could be effective for separation of *N. sativa* seeds from impurities. The overall accuracy of the suggested ANN model for the online classification of the *N. sativa* class was 97.04%. The impurities were distinguished with average sensitivity of 95.57%. The proposed algorithm

accurately separated the *N. sativa* class from its impurity. For industrial applications, the number of grain flow lines could be increased to reach an acceptable capacity. In addition, high-speed pneumatic or electromechanical actuators could be used for adaptation of the suggested approach in industrial applications. The developed methodology can also be used for non-destructive separation of other similar seeds.

Authors' contributions

Conceptualization: S. Alvandi, S. S. Mohtasebi, M. Omid

Data curation: S. Alvandi, S. S. Mohtasebi

Formal analysis: S. Alvandi, S. S. Mohtasebi, M. Omid, M. Hosseinpour-Zarnaq

Funding acquisition: S. S. Mohtasebi

Investigation: S. Alvandi, S. S. Mohtasebi

Methodology: S. Alvandi, S. S. Mohtasebi

Project administration: S. S. Mohtasebi

Resources: S. Alvandi, M. Hosseinpour-Zarnaq

Software: S. Alvandi, S. S. Mohtasebi, M. Omid

Supervision: S. S. Mohtasebi

Validation: S. S. Mohtasebi, M. Omid

Visualization: S. S. Mohtasebi

Writing – original draft: S. Alvandi, S. S. Mohtasebi, M. Omid, M. Hosseinpour-Zarnaq

Writing – review & editing: S. S. Mohtasebi, M. Omid, M. Hosseinpour-Zarnaq

References

- Bao Y, Mi C, Wu N, Liu F, He Y, 2019. Rapid classification of wheat grain varieties using hyperspectral imaging and chemometrics. *Appl Sci* 9(19): 4119. <https://doi.org/10.3390/app9194119>
- Banakar A, Zareiforush H, Baigvand M, Montazeri M, Khodaei J, Behroozi-Khazaei N, 2017. Combined application of decision tree and fuzzy logic techniques for intelligent grading of dried figs. *J Food Process Eng* 40(3): e12456. <https://doi.org/10.1111/jfpe.12456>
- Copeland LO, McDonald MF, 2012. Principles of seed science and technology. Springer Sci & Bus Media.
- Ebrahimi E, Mollazade K, Babaei S, 2014. Toward an automatic wheat purity measuring device: A machine vision-based neural networks-assisted imperialist competitive algorithm approach. *Measurement* 55: 196-205. <https://doi.org/10.1016/j.measurement.2014.05.003>
- Fayyazi S, Abbaspour-Fard MH, Rohani A, Monadjemi SA, Sadriani H, 2017. Identification and classification of three Iranian rice varieties in mixed bulks using image processing and MLP neural network. *Int J Food Eng* 13(5): 20160121. <https://doi.org/10.1515/ijfe-2016-0121>
- Gharby S, Harhar H, Guillaume D, Roudani A, Boulbaroud S, Ibrahim M, et al., 2015. Chemical investigation of

- Nigella sativa* L. seed oil produced in Morocco. *J Saudi Soc Agric Sci* 14(2): 172-177. <https://doi.org/10.1016/j.jssas.2013.12.001>
- Granitto PM, Verdes PF, Ceccatto HA, 2005. Large-scale investigation of weed seed identification by machine vision. *Comput Electron Agric* 47(1): 15-24. <https://doi.org/10.1016/j.compag.2004.10.003>
- Han XZ, Wang KJ, Yuan Y, Chen C, Liang L, 2014. Research on grading detection of the wheat seeds. *Sci World J* 2014: 702906. <https://doi.org/10.1155/2014/702906>
- Jolliffe IT, Cadima J, 2016. Principal component analysis: a review and recent developments. *Philos T Roy Soc A* 374(2016): 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Kanjanawanishkul K, Chupawa P, Nuantoon T, 2018. Design and assessment of an automated sweet pepper seed sorting machine. *Eng Agr Environ Food* 11(4): 196-201. <https://doi.org/10.1016/j.eaef.2018.05.001>
- Liu D, Ning X, Li Z, Yang D, Li H, Gao L, 2015. Discriminating and elimination of damaged soybean seeds based on image characteristics. *Stored Prod Res* 60: 67-74. <https://doi.org/10.1016/j.jspr.2014.10.001>
- Madhavan J, Salim M, Durairaj U, Kotteeswaran R, 2020. Wheat seed classification using neural network pattern recognizer. *Mater Today Proc*. <https://doi.org/10.1016/j.matpr.2021.03.226>
- Majdalawieh AF, Fayyad MW, 2016. Recent advances on the anti-cancer properties of *Nigella sativa*, a widely used food additive. *J Ayurveda Integr Med* 7(3): 173-180. <https://doi.org/10.1016/j.jaim.2016.07.004>
- Mollazade K, Omid M, Arefi A, 2012a. Comparing data mining classifiers for grading raisins based on visual features. *Comput Electron Agric* 84: 124-131. <https://doi.org/10.1016/j.compag.2012.03.004>
- Mollazade K, Omid M, Tab FA, Mohtasebi SS, 2012b. Principles and applications of light backscattering imaging in quality evaluation of agro-food products: A review. *Food Bioprocess Technol* 5(5): 1465-1485. <https://doi.org/10.1007/s11947-012-0821-x>
- Mollazade K, 2017. Non-destructive identifying level of browning development in button mushroom (*Agaricus bisporus*) using hyperspectral imaging associated with chemometrics. *Food Anal Meth* 10(8): 2743-2754. <https://doi.org/10.1007/s12161-017-0845-y>
- Nouri-Ahmadabadi H, Omid M, Mohtasebi SS, Firouz MS, 2017. Design, development and evaluation of an online grading system for peeled pistachios equipped with machine vision technology and support vector machine. *Inf Process Agric* 4(4): 333-341. <https://doi.org/10.1016/j.inpa.2017.06.002>
- Orlandi G, Calvini R, Foca G, Ulrici A, 2018. Automated quantification of defective maize kernels by means of multivariate image analysis. *Food Control* 85: 259-268. <https://doi.org/10.1016/j.foodcont.2017.10.008>
- Pearson T, Moore D, Pearson J, 2012. A machine vision system for high speed sorting of small spots on grains. *Food Measure* 6(1): 27-34. <https://doi.org/10.1007/s11694-012-9130-3>
- Qadri S, Furqan Qadri S, Razzaq A, Ul Rehman M, Ahmad N, Nawaz SA, et al., 2021. Classification of canola seed varieties based on multi-feature analysis using computer vision approach. *Int J Food Prop* 24(1): 493-504. <https://doi.org/10.1080/10942912.2021.1900235>
- Quan L, Zhang T, Sun L, Chen X, Xu Z, 2018. Design and testing of an on-line omnidirectional inspection and sorting system for soybean seeds. *Appl Eng Agric* 34(6): 1003-1016. <https://doi.org/10.13031/aea.12935>
- Rad SJ, Tab FA, Mollazade K, 2011. Classification of rice varieties using optimal color and texture features and BP neural networks. 7th Iran Conf on Machine Vision and Image Processing, Nov 16, pp. 1-5. IEEE. <https://doi.org/10.1109/IranianMVIP.2011.6121583>
- Rahman A, Cho BK, 2016. Assessment of seed quality using non-destructive measurement techniques: A review. *Seed Sci Res* 26(4): 285-305. <https://doi.org/10.1017/S0960258516000234>
- Shen W, Cheng F, Zhang Y, Wei X, Fu Q, Zhang Y, 2020. Automatic recognition of ingestive-related behaviors of dairy cows based on triaxial acceleration. *Inf Process Agric* 7(3): 427-443. <https://doi.org/10.1016/j.inpa.2019.10.004>
- Shrestha BL, Kang YM, Yu D, Baik OD, 2016. A two-camera machine vision approach to separating and identifying laboratory sprouted wheat kernels. *Biosyst Eng* 147: 265-273. <https://doi.org/10.1016/j.biosystemseng.2016.04.008>
- Stanimirova I, Daszykowski M, Walczak B, 2013. Robust methods in analysis of multivariate food chemistry data. *Data Handling Sci Technol* 1(28): 315-340. <https://doi.org/10.1016/B978-0-444-59528-7.00008-9>
- Teimouri N, Omid M, Mollazade K, Rajabipour A, 2014. A novel artificial neural networks assisted segmentation algorithm for discriminating almond nut and shell from background and shadow. *Comput Electron Agric* 105: 34-43. <https://doi.org/10.1016/j.compag.2014.04.008>
- Teimouri N, Omid M, Mollazade K, Mousazadeh H, Alimardani R, Karstoft H, 2018. On-line separation and sorting of chicken portions using a robust vision-based intelligent modelling approach. *Biosyst Eng* 167: 8-20. <https://doi.org/10.1016/j.biosystemseng.2017.12.009>
- Valiente-Gonzalez JM, Andreu-Garcia G, Potter P, Rodas-Jorda A, 2014. Automatic corn (*Zea mays*) kernel inspection system using novelty detection based on principal component analysis. *Biosyst Eng* 117: 94-103. <https://doi.org/10.1016/j.biosystemseng.2013.09.003>
- Vithu P, Moses JA, 2016. Machine vision system for food grain quality evaluation: A review. *Trends Food Sci Technol* 1(56): 13-20. <https://doi.org/10.1016/j.tifs.2016.07.011>
- Wakholi C, Kandpal LM, Lee H, Bae H, Park E, Kim MS, et al., 2018. Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics. *Sens Actuators B Chem* 255: 498-507. <https://doi.org/10.1016/j.snb.2017.08.036>